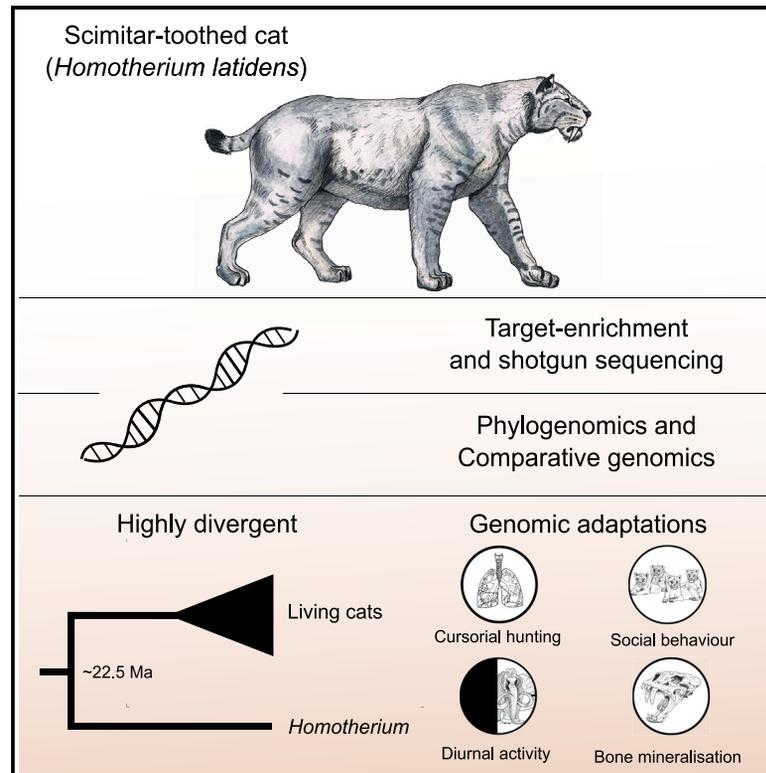


Current Biology

Genomic Adaptations and Evolutionary History of the Extinct Scimitar-Toothed Cat, *Homotherium latidens*

Graphical Abstract



Authors

Ross Barnett, Michael V. Westbury, Marcela Sandoval-Velasco, ..., Jong Bhak, Nobuyuki Yamaguchi, M. Thomas P. Gilbert

Correspondence

m.westbury@sund.ku.dk (M.V.W.),
tgilbert@sund.ku.dk (M.T.P.G.)

In Brief

Here, Barnett et al. sequence the nuclear genome of *Homotherium latidens* through a combination of shotgun and target-capture approaches. Analyses confirm *Homotherium* to be a highly divergent lineage from all living cat species (~22.5 Ma) and reveal genes under selection putatively related to a cursorial and diurnal hunting behavior.

Highlights

- Nuclear genome and exome analyses of extinct scimitar-toothed cat, *Homotherium latidens*
- *Homotherium* was a highly divergent lineage from all living cat species (~22.5 Ma)
- Genetic adaptations to cursorial and diurnal hunting behaviors
- Relatively high levels of genetic diversity in this individual



Report

Genomic Adaptations and Evolutionary History of the Extinct Scimitar-Toothed Cat, *Homotherium latidens*

Ross Barnett,^{1,35} Michael V. Westbury,^{1,35,36,*} Marcela Sandoval-Velasco,^{1,35} Filipe Garrett Vieira,¹ Sungwon Jeon,^{2,3} Grant Zazula,⁴ Michael D. Martin,⁵ Simon Y.W. Ho,⁶ Niklas Mather,⁶ Shyam Gopalakrishnan,^{1,7} Jazmín Ramos-Madrigal,^{1,7} Marc de Manuel,⁸ M. Lisandra Zepeda-Mendoza,^{1,9} Agostinho Antunes,^{10,11} Aldo Carmona Baez,¹ Binia De Cahsan,¹ Greger Larson,¹² Stephen J. O'Brien,^{13,14} Eduardo Eizirik,^{15,32,33}

(Author list continued on next page)

¹Section for Evolutionary Genomics, The GLOBE Institute, University of Copenhagen, Øster Voldgade 5-7, Copenhagen, Denmark

²Korean Genomics Center (KOGIC), Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea

³Department of Biomedical Engineering, School of Life Sciences, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea

⁴Yukon Palaeontology Program, Department of Tourism and Culture, Government of Yukon, PO Box 2703, Whitehorse, YT Y1A 2C6, Canada

⁵Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology (NTNU), Trondheim NO-7491, Norway

⁶School of Life and Environmental Sciences, University of Sydney, NSW 2006, Australia

⁷Center for Evolutionary Hologenomics, The GLOBE Institute, University of Copenhagen, Øster Farimagsgade 5A, Copenhagen 1352, Denmark

⁸Institute of Evolutionary Biology (UPF-CSIC), PRBB, Dr. Aiguader 88, Barcelona 08003, Spain

⁹School of Medical and Dental Sciences, Institute of Microbiology and Infection, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

¹⁰CIIMAR/CIMAR, Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Terminal de Cruzeiros do Porto de Leixões, Av. General Norton de Matos, s/n, Porto 4450-208, Portugal

¹¹Department of Biology, Faculty of Sciences, University of Porto, Porto 4169-007, Portugal

(Affiliations continued on next page)

SUMMARY

Homotherium was a genus of large-bodied scimitar-toothed cats, morphologically distinct from any extant felid species, that went extinct at the end of the Pleistocene [1–4]. They possessed large, saber-form serrated canine teeth, powerful forelimbs, a sloping back, and an enlarged optic bulb, all of which were key characteristics for predation on Pleistocene megafauna [5]. Previous mitochondrial DNA phylogenies suggested that it was a highly divergent sister lineage to all extant cat species [6–8]. However, mitochondrial phylogenies can be misled by hybridization [9], incomplete lineage sorting (ILS), or sex-biased dispersal patterns [10], which might be especially relevant for *Homotherium* since widespread mito-nuclear discrepancies have been uncovered in modern cats [10]. To examine the evolutionary history of *Homotherium*, we generated a ~7x nuclear genome and a ~38x exome from *H. latidens* using shotgun and target-capture sequencing approaches. Phylogenetic analyses reveal *Homotherium* as highly divergent (~22.5 Ma) from living cat species, with no detectable signs of gene flow. Comparative genomic analyses found signatures of positive selection in several genes, including those involved in vision, cognitive function, and energy consumption, putatively consistent with diurnal activity, well-developed social behavior, and cursorial hunting [5]. Finally, we uncover relatively high levels of genetic diversity, suggesting that *Homotherium* may have been more abundant than the limited fossil record suggests [3, 4, 11–14]. Our findings complement and extend previous inferences from both the fossil record and initial molecular studies, enhancing our understanding of the evolution and ecology of this remarkable lineage.

RESULTS AND DISCUSSION

We used a combination of shotgun and whole genome and exome target-capture Illumina sequencing to generate the nuclear genome of a single *Homotherium latidens* individual to a depth

of ~7x coverage, and its exome to ~38x coverage. The genome was sequenced from a fossil humerus (specimen YG 439.38), which was determined to be older than the limits of radiocarbon dating (>47.5 kya [UCIAMS-142835]), recovered from Pleistocene permafrost sediments near Dawson City, Yukon Territory, Canada



Warren E. Johnson,^{16,17,18} Klaus-Peter Koepfli,¹⁶ Andreas Wilting,¹⁹ Jörns Fickel,^{19,20} Love Dalén,^{21,22} Eline D. Lorenzen,^{1,7} Tomas Marques-Bonet,^{8,23,24,25} Anders J. Hansen,^{7,34} Guojie Zhang,^{26,27,28} Jong Bhak,^{2,3,29,30} Nobuyuki Yamaguchi,³¹ and M. Thomas P. Gilbert^{1,5,7,*}

¹²The Palaeogenomics and Bio-Archaeology Research Network, Research Laboratory for Archaeology and History of Art, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK

¹³Laboratory of Genomic Diversity, Center for Computer Technologies, ITMO University, 49 Kronverkskiy Pr., St. Petersburg 197101, Russia

¹⁴Guy Harvey Oceanographic Center, Halmos College of Natural Sciences and Oceanography, Nova Southeastern University, 8000 North Ocean Drive, Ft Lauderdale, FL 33004, USA

¹⁵Laboratory of Genomics and Molecular Biology, Escola de Ciências da Saúde e da Vida, PUCRS, Porto Alegre, RS, Brazil

¹⁶Center for Species Survival, Smithsonian Conservation Biology Institute, National Zoological Park, 1500 Remount Road, Front Royal, VA 22630, USA

¹⁷The Walter Reed Biosystematics Unit, Museum Support Center MRC-534, Smithsonian Institution, 4210 Silver Hill Rd., Suitland, MD 20746-2863, USA

¹⁸Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA

¹⁹Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Str. 17, Berlin 10315, Germany

²⁰Institute for Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Strasse 24-25, Potsdam 14476, Germany

²¹Centre for Palaeogenetics, Svante Arrhenius väg 20C, Stockholm SE-10691, Sweden

²²Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, Stockholm 10405, Sweden

²³CNAG-CRG, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Baldri Reixac 4, Barcelona 08028, Spain

²⁴Institució Catalana de Recerca i Estudis Avançats, ICREA, Barcelona 08003, Spain

²⁵Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, c/ Columnes s/n, Cerdanyola del Vallès, Barcelona 08193, Spain

²⁶BGI-Shenzhen, Shenzhen 518083, China

²⁷Section for Ecology and Evolution, Department of Biology, Faculty of Science, University of Copenhagen, Universitetsparken 15, Copenhagen, Denmark

²⁸State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China

²⁹Clinomics, Inc., Ulsan 44919, Republic of Korea

³⁰Personal Genomics Institute (PGI), Genome Research Foundation (GRF), Osong 28160, Republic of Korea

³¹Institute of Tropical Biodiversity and Sustainable Development, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu 21030, Malaysia

³²INCT Ecologia, Evolução e Conservação da Biodiversidade (INCT-EECBio), Goiânia, GO, Brazil

³³Instituto Pró-Carnívoros, Atibaia, SP, Brazil

³⁴Section for GeoGenetics, The GLOBE Institute, University of Copenhagen, Øster Voldgade 5-7, Copenhagen, Denmark

³⁵These authors contributed equally

³⁶Lead Contact

*Correspondence: m.westbury@sund.ku.dk (M.V.W.), tgilbert@sund.ku.dk (M.T.P.G.)

<https://doi.org/10.1016/j.cub.2020.09.051>

[8]. This nuclear dataset represents a large increase in the number of independent genetic loci available for analysis, each of which has been uniquely passed down over many generations of lineage sorting, drift, selection, and recombination. By investigating such a large number of independent loci, we are able to better understand the *Homotherium* lineage and make broad evolutionary inferences that are not possible with datasets comprising a more limited number of markers. Therefore, we refer to the genus without mentioning the species name throughout the manuscript when results can be interpreted as representing the lineage as a whole. After filtering for quality and mapping against the lion (*Panthera leo*) nuclear genome [15] (NCBI accession: JAAVKH000000000), we undertook phylogenomic analyses using a supermatrix of the exonic data, as well as a combined tree constructed from multiple individual gene trees. The inferred topologies are consistent with those of previous studies based on both fossil [2] and mitochondrial datasets [8] in which *Homotherium* is a sister lineage to all extant cat species (Figures 1 and S1).

We next dated the divergence between *Homotherium* and the extant cat species using MCMCTree [16] and four fossil calibrations (see STAR Methods and Table S1). We found that *Homotherium* diverged from extant cat species ~22.5 Ma (95%

credibility interval [CI] 17.4–27.8 Ma), close to the Oligocene/Miocene boundary ~23 Ma. This timing coincides with either late *Proailurus* sp. or early *Pseudaelurus* sp. as potential ancestors [17]. Our inferred date is similar to the estimate of ~20 Ma (95% CI 18.2–22.0 Ma) that was previously determined based on mitochondrial genomes [8]. This deep divergence confirms that *Homotherium* belonged to a distinct clade from all living felids, providing support for Machairodontinae as a separate subfamily from Felinae.

Still, the clear-cut phylogenetic relationships reported here might not represent the full evolutionary relationship of *Homotherium* to living felids. Several recent genomic studies have demonstrated that interspecific hybridization is much more common than previously believed in several lineages within Carnivora, including crown canids [18], hyenas [9], and extant cats [10, 19, 20]. For this reason, we explored our dataset for evidence of topological incongruencies that could arise due to incomplete lineage sorting (ILS) or potentially indicate gene flow. In regard to gene flow, *Homotherium* is of considerable interest, as it once spanned from southern Africa, across Eurasia and North America, to South America, arguably the largest geographical range of all the saber-toothed cats

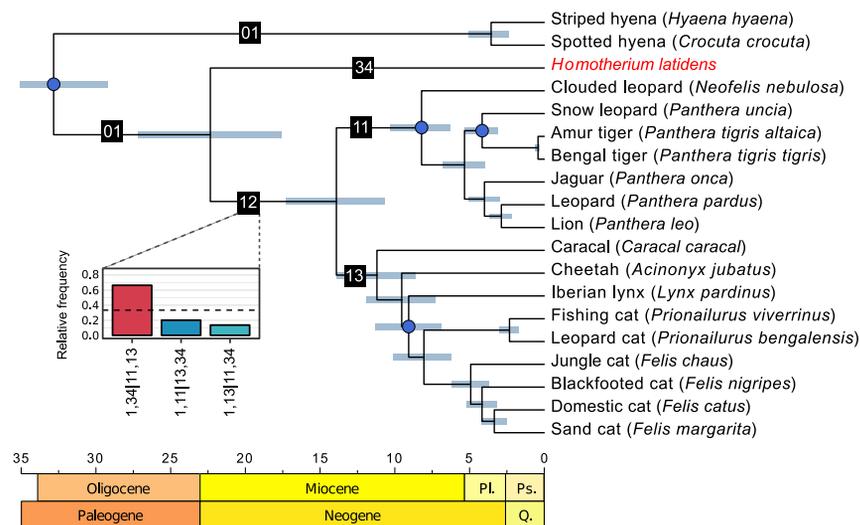


Figure 1. Evolutionary Time-Tree of 17 Cats and Two Hyenas

Tree topology inferred using both RAxML and ASTRAL-II, with node ages estimated using a Bayesian relaxed-clock analysis of concatenated sequences of 21,034 exons (total length 29,216,712 bp) (Figure S1). A separate correlated-rates relaxed clock was applied to each of the three codon positions. Relative frequencies of the three possible bipartitions (possible arrangements of a quartet on an unrooted tree) are shown for the internal branch containing *Homotherium* (branch 12) and the remaining nodes in Figure S2. Dashed lines show the threshold value of one-third, shown theoretically to be the minimum frequency for a true bipartition. Relevant branch labels have been given based on those in Figure S2. Blue horizontal bars represent 95% credibility intervals of node times. Blue circles indicate internal nodes with fossil-based age constraints. A geological timescale is shown below the tree (Q. = Quaternary, Pl. = Pliocene, Ps. = Pleistocene). Ages of key nodes in the phylogeny can be found in Table S1.

(Machairodontinae), although it is uncertain whether it occurred in all those regions simultaneously [2–4, 14, 21]. In addition to occupying a huge geographical distribution, it also colonized a variety of different habitats, from densely vegetated Java to open steppe-tundra across the Holarctic [13, 22]. Furthermore, fossil evidence suggests that *Homotherium* succeeded in expanding its distribution despite potential competition with other sympatric large cats, including lions (*Panthera leo* and the now-extinct cave lion *P. spelaea*), leopards (*P. pardus*) and the now-extinct dirk-toothed cat *Megantereon cultridens* across Eurasia and Africa, tigers (*P. tigris*) in Southeast Asia, and jaguars (*P. onca*), American lions (*P. atrox*), and other now-extinct saber-toothed cats (*Smilodon* spp.) in the Americas [2, 23].

To explore whether gene flow may have occurred between *Homotherium* and living cat species, we investigated conflicting phylogenetic signals in our phylogenomic tree using ASTRAL-II [24] and DiscoVista [25] (Figures 1 and S2). Although our results reproduce some of the notorious cases of genealogical discordance in other felids (e.g., the relationships among lion, jaguar, and leopard [node 6]; or the relative placement of the *Lynx* and *Puma* lineages [node 15]), gene-tree quartet frequencies of alternative topologies for the node separating *Homotherium* from all other cat species were lower than would be expected (less than one-third) if gene flow or ILS had played an important role in this divergence (Figure 1). This implies no gene flow between the lineage that our *Homotherium* sample represents and the extant cat species, but does not exclude the possibility of undetectable gene flow between the ancestor of all extant cat lineages and the *Homotherium* lineage. The most plausible explanation for the inferred lack of gene flow is the deep divergence between *Homotherium* and extant cat species. The oldest signal of gene flow we would be able to detect with the current dataset would be after the divergence of stem Felinae ~14 Ma (Figure 1). This means that any gene flow over an ~8-million-year period between Machairodontinae and Felinae is undetectable. In contrast, the radiation of the main lineages of modern cats occurred within a ~5-million-year period (Figure 1). This rapid radiation may have

allowed gene flow to occur between these lineages before they became too genetically divergent from one another to produce viable offspring [10]. Although alternative explanations could apply to the apparent lack of gene flow, these are much more unlikely. One possibility is that *Homotherium* were simply unable to interact with other cat species due to either eco-geographical barriers, competitive exclusion, or low population density. The former seems implausible given their above-described distribution and affinities to very different habitats. Competitive exclusion also seems unlikely, as *Homotherium* fossils are found at the same sites as *Panthera*, and they are likely contemporaneous [23]. Low population density could be interpreted as more likely, as the fossil record of *Homotherium* is considerably more fragmentary than those of other contemporary large cats, including *Smilodon* and *Panthera* (especially lions), tentatively suggesting that *Homotherium* occurred at lower population densities [3, 4, 11–14]. However, even at low densities, occasional contact with sympatric species would not be precluded.

Another alternative explanation is that behavioral and/or other ecological mechanisms may have prevented hybridization. An example of this can be seen between extant lions and leopards. Though they often occupy the same area, leopards are known to actively avoid lions [26]. A similar behavioral and/or ecological mechanism of avoidance may have occurred between *Homotherium* and other sympatric felids. To further add evidence to our finding of a lack of gene flow, although archaic admixture between living mammalian species and unknown, unsampled lineages has been uncovered (e.g., humans [27], bonobos [28], and dogs [18]), including signatures of introgression into leopards from an extinct Felinae species closely related to *Panthera* [19], no signs of archaic admixture between living cat species and an unknown lineage as divergent as *Homotherium* have been detected in previous studies.

This divergence thus leads to the question of what genetic adaptations occurred as these lineages differentiated from one another. To address the question of what genomic underpinnings made *Homotherium* unique, we performed a comparative

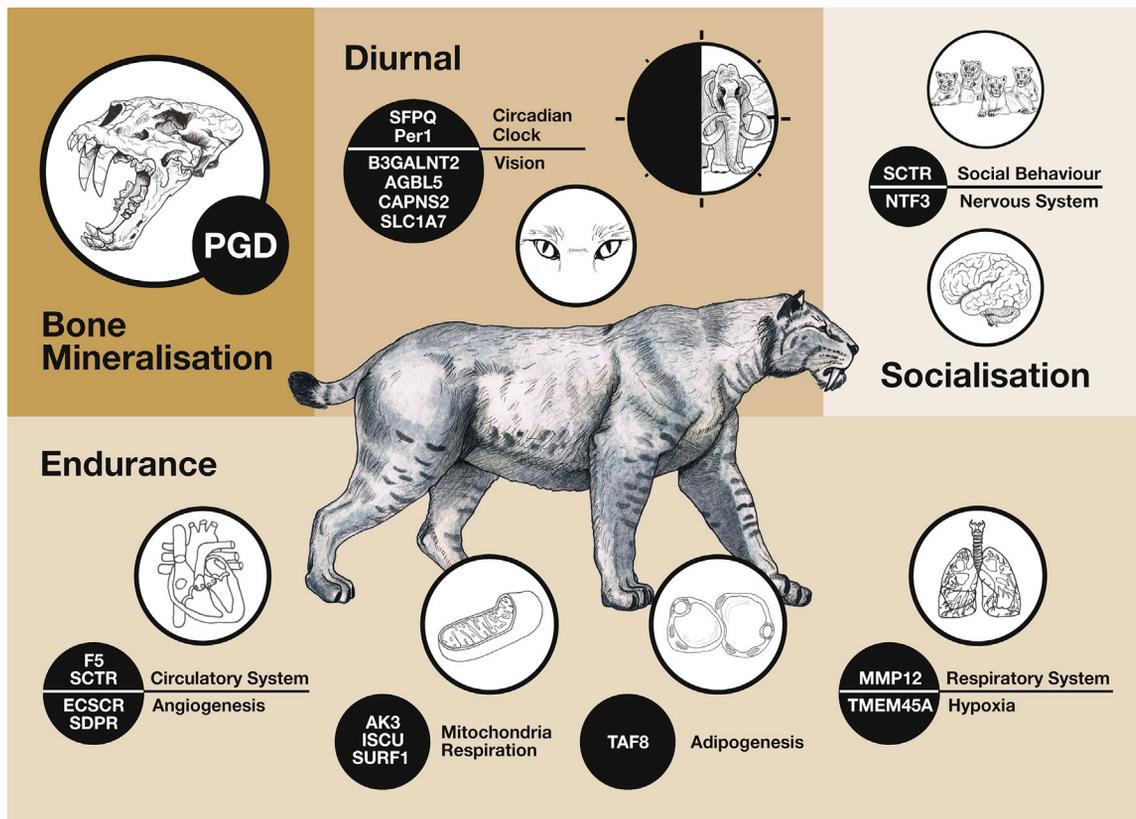


Figure 2. Depiction of 18 of the 31 Genes under Positive Selection with High Values (Free Ratio > 2) in the *Homotherium* Genome

Hypothetical functions and the adaptive insights that these provide on the species' behavior, morphology, and functional adaptations are also shown. Additional genes not depicted here are likely involved in cellular processes such as apoptosis, protein synthesis, and protein signaling, as well as immunity/cancer, olfaction, and reproduction (Table S3). All genes showing significant signs of positive selection can be seen in Table S2.

genomics analysis and uncovered signatures of positive selection in several protein-coding regions across its genome (Table S2). We found evidence of positive selection ($5 > d_N/d_S$ ratio > 1) in 230 genes out of the 2,191 assessed 1:1 orthologous loci. Of these 230 genes, 31 were deemed to be highly significant (a free ratio > 2) and were further explored for putative functions and phenotypic roles (Table S3). Several highly significant positively selected genes were consistent with a putative diurnal behavior in *Homotherium* [5] (Figure 2). Strong positive selection was detected in genes linked to vision (B3GALNT2, AGBL5, CAPNS2, and SLC1A7) that have known phenotypes involving retinal degeneration, retinitis pigmentosa, hydrolysis of natural lens proteins, and visual processing, e.g., [29, 30]. We also found evidence of positive selection in genes linked with entrainment and the master regulation of circadian clock rhythm (SFPQ and Per1) [31, 32]. Although not in the same genes specifically, previous studies have shown polymorphisms in circadian clock regulatory genes to be related to diurnal preference [33, 34], strengthening the link between circadian genes and diurnal behavior. While speculative, these results do provide some support toward the idea that *Homotherium* hunted during the day, unlike many extant cat species, which are either crepuscular or nocturnal [35]. This hypothesis is further supported by several anatomical features of this species, including an enlarged optic bulb and a large and complex visual cortex [5].

Signals of positive selection were also inferred in genes that were plausibly involved in adaptations helping to increase endurance for a cursorial hunting style in *Homotherium*. These include genes with major influences on the respiratory system/hypoxia (TMEM45A, which has been tied to the deprivation of oxygen supply), the circulatory system (F5 and MMP12), angiogenesis (ECSCR—involved in the formation of new blood vessels), adipogenesis (TAF8—a positive adipogenesis regulator playing a vital role in energy homeostasis), the respiratory/circulatory systems (MMP12, with a role in aneurysm, a result of a weakened blood vessel wall), and mitochondrial respiration (AK3, ISCU, and SURF, with relevance in skeletal muscle and the myopathy phenotype with severe exercise intolerance, e.g., [36] and the biogenesis of the cytochrome c oxidase [COX] complex). Novel adaptations in these genes may have enabled sustained running necessary for hunting in more open habitats and the pursuit of prey until their point of exhaustion. The synergistic interplay of these various functional enhancements could have also been assisted by improvements in bone mineralization (PGD). Improved bone mineralization would have been invaluable to develop and sustain a robust skeletal framework and powerful forelimbs needed for a putatively cursorial hunting style. Furthermore, although not the PGD gene specifically, two other genes playing a role in bone development and repair (DMP1 and PTN) have been found to be under positive selection in a number of

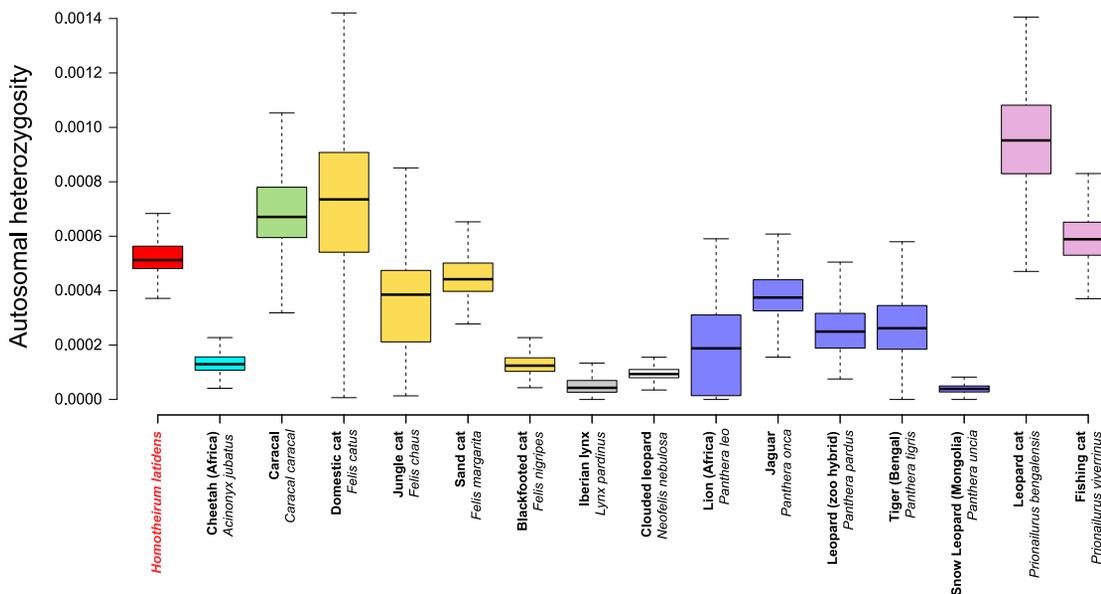


Figure 3. Autosomal Heterozygosity Estimates for Each Species Included in the Current Study

Variance was estimated by calculating the average heterozygosity for each scaffold independently. Colors represent the genus that each individual belongs to (red: *Homotherium*; cyan: *Acinonyx*; green: *Caracal*; yellow: *Felis*; gray: *Lynx*; white: *Neofelis*; blue: *Panthera*; pink: *Prionailurus*). Exome-wide heterozygosity can be seen in Figure S3.

carnivore genomes (including several felid species) [37], suggesting that robust bones may be key for adaptation to predatory behavior. The signatures of selection in the above-mentioned genes add additional evidence to a cursorial hunting style suggested by post-cranial skeletal morphological data, including less-retractable claws, which are thought to improve traction during long-distance, medium-speed travel, similar to canids or hyenas [23], along with a higher brachial index (radius to humerus ratio) [2, 14, 38, 39].

Finally, although social behavior is very complex and therefore difficult to directly link to certain genetic traits, we found evidence of positive selection in genes that putatively play roles in cognition/behavior (SCTR, which is linked to impaired synaptic plasticity and social behavior [Mouse Genome Informatics—MGI database—<http://www.informatics.jax.org>]) and the nervous system (NTF3—influencing the nerve growth factor). We speculate that these genes may have played a role in coordinated social interactions that others have suggested were required for *Homotherium* to successfully hunt large prey species [13, 14], but additional data would be needed to support this notion. Although our results do not provide definitive links between genes under positive selection and their ecological consequences, we provide the starting ground for functional studies into the relationship between the genetics and ecology of the extinct hypercarnivore *Homotherium*. We find a number of positively selected genes whose known functions complement the paleontological data and represent the putative genetic underpinnings for several unique adaptations of *Homotherium* that appear to be closely linked to cursorial group-hunting strategies and diurnal behavior.

As mentioned above, the fossil record for *Homotherium* is considerably more fragmentary than those of other contemporary

large cats, including *Smilodon* and *Panthera* (especially lions), tentatively suggesting that *Homotherium* occurred at lower population densities [3, 4, 11–14]. This leads to the question of whether the fragmentary fossil record of *Homotherium* truly reflects low population densities or is a stochastic result based on the likelihood of preservation due to their habitat or behavior. To examine the relative abundance of *Homotherium latidens* compared with extant cat species, we compared the genetic diversity (which is correlated with the effective population size), in the form of heterozygosity, from our *Homotherium latidens* genome to that of single individual representatives from 15 extant cat species included in this study. We used two different approaches (autosomal-wide and exome-wide) to estimate heterozygosity. To ensure comparability between species when calculating heterozygosity, we only considered transversions to remove biases that could occur due to the much higher ancient DNA (aDNA) damage levels expected in our *Homotherium* individual relative to the extant species. For the autosomal-wide heterozygosity estimates, to further ensure comparability between results, we downsampled all individuals to a comparable read depth to the *Homotherium latidens* genome before estimating heterozygosity. As we performed exome capture on our *Homotherium latidens* individual and therefore had relatively high coverage, we did not perform any downsampling with this dataset. We caution that our estimations are based on estimates from single representatives of each species, and thus may not hold true species-wide. However, it is striking that our *Homotherium latidens* genome exhibits medium to high levels of genetic diversity, higher than that inferred for the other big cat species. This inference held true regardless of whether autosomal-wide (Figure 3) or exome-wide heterozygosity (Figure S3) was compared. Given the relationship between genome-wide heterozygosity, effective population size, and census size, this finding suggests that

Homotherium latidens was relatively abundant in contrast to previously suggested low population densities based on low levels of mitochondrial genetic diversity compared with other species of large cats [8]. However, as this previous suggestion was purely based on the single-locus mitochondrial genome from three individuals, the inference on diversity was limited.

Our finding of a putative high abundance in *Homotherium latidens* is strengthened by *Homotherium* having had the largest geographical range of all Machairodontinae, ranging from southern Africa, across Eurasia and North America, to South America [2–4, 14, 21]. Although genetic diversity often reflects population size, the demography and life history traits of the population/species can also play a role. One notable factor that may have resulted in higher levels of genetic diversity in *Homotherium* occupying the (sub)Arctic would be long-distance migration between structured populations. In general, species (including cursorial predators) dwelling in areas with dramatic seasonal variability or low primary productivity (e.g., the [sub]Arctic) travel greater distances than those in more stable and productive regions [40]. Therefore, it may have been possible that *Homotherium*, as a cursorial predator living in the less-productive (sub)Arctic region, moved across an enormous range, leading to increased gene flow between distant populations. Regardless, the huge geographical distribution of *Homotherium*, and the apparent ability of the genus to colonize different habitats, suggests it was a very successful taxon and fits our inference that it may have been relatively abundant.

The apparent success of *Homotherium* leads to the question of why the lineage was unable to survive to the present. Although it cannot be known for certain, some of the exact adaptations/specializations that led to *Homotherium*'s success could also have led to its downfall. Toward the end of the Late Pleistocene, a decrease in large prey availability may have caused more direct competition with other cat species that were likely more effective at capturing the remaining smaller prey species. The specific adaptations *Homotherium* had acquired would have suddenly become less advantageous, leaving the lineage stuck on a path to extinction.

Conclusions

The genome sequence of the scimitar-toothed cat, *Homotherium latidens*, improves our understanding of its evolutionary relationships with other extant cat species and into the genetic underpinnings of its unique adaptations. Our results demonstrate that *Homotherium* was highly divergent (~22.5 Ma) from all living cat species and did not undergo any detectable gene flow with extant felid species after their initial radiation ~14 Ma. This placement supports the recognition of Machairodontinae as a distinct subfamily within the Felidae. Furthermore, we found evidence of positive selection in several genes involved in vision, cognitive function, and energy consumption, potentially consistent with the diurnal and hunting/social behavior of this extinct lineage. Finally, we uncovered relatively high levels of genetic diversity in our *Homotherium latidens* individual, suggesting that it was not only a successful lineage, but also rather abundant relative to extant cat species. Our study shows how the fossil record and paleogenomics can be used synergistically to better

understand the evolution and ecology of extinct species that lack extant close relatives for comparison.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead Contact
 - Materials Availability
 - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Fossil information
 - Radiocarbon dating
- METHOD DETAILS
 - Ancient DNA extraction, library preparation, and sequencing
 - Genome capture
 - Modern DNA extraction, library preparation, and sequencing
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Data processing pipeline
 - Extracting coding and protein sequences
 - Phylogenetic analyses
 - Phylogenomic dating
 - Tests of positive selection
 - Genetic diversity

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cub.2020.09.051>.

ACKNOWLEDGMENTS

This project received funding from the European Union's Seventh Framework Programme for Research, Technological Development, and Demonstration under grant agreement no. FP7-PEOPLE-2011-IEF-298820 and ERC Consolidator Award 681396—Extinction Genomics to M.T.P.G. Portions of this manuscript were prepared while W.E.J. held a National Research Council Research Associateship Award at the Walter Reed Army Institute of Research (WRAIR). The material has been reviewed by WRAIR and there is no objection to its presentation and/or publication. The opinions and assertions contained herein are the private views of the authors, and are not to be construed as official or as reflecting true views of the Department of the Army or the Department of Defense. We thank the laboratory technicians of the Centre for GeoGenetics and the staff of the Danish National High-Throughput DNA Sequencing Centre for technical assistance. We thank Tom Stafford, Jr. and Stafford Research, LLC for radiocarbon dating and discussion. Thanks to the field crew that recovered bone YG 439.38: Beth Shapiro, Matthias Stiller, Duane Froese, and Tyler Kuhn. We are grateful to the placer gold mining community and the Tr'ondëk Hwëch'in for their continued support and partnership with our research in the Klondike goldfields region. We thank Rheon Slade for his assistance on the layout of Figure 2. Lastly, we thank Jean-Marie Rouillard (Arbor Biosciences) for help with bait design.

AUTHOR CONTRIBUTIONS

Conceptualization, R.B., N.Y., and M.T.P.G.; Formal Analysis, M.V.W., F.G.V., S.J., M.D.M., S.Y.W.H., N.M., S.G., J.R.-M., M.d.M., M.L.Z.-M., A.A., and A.C.B.; Investigation, R.B. and M.S.-V.; Writing – Original Draft, R.B.,

M.V.W., M.S.-V., N.Y., and M.T.P.G.; Writing – Review & Editing, all authors; Visualization, M.V.W., S.Y.W.H., S.G., A.A., and B.D.C.; Funding Acquisition, R.B. and M.T.P.G.; Resources, G.Z., G.L., S.J.O., E.E., W.E.J., K.-P.K., A.W., J.F., L.D., E.D.L., T.M.-B., A.H., G.Z., J.B., N.Y., and M.T.P.G.; Supervision, M.T.P.G., T.M.-B., and J.B.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 28, 2020

Revised: July 10, 2020

Accepted: September 15, 2020

Published: October 15, 2020

REFERENCES

- Reumer, J.W.F., Rook, L., Van Der Borg, K., Post, K., Mol, D., and De Vos, J. (2003). Late Pleistocene survival of the saber-toothed cat *Homotherium* in northwestern Europe. *J. Vertebr. Paleontol.* *23*, 260–262.
- Turner, A., and Antón, M. (1997). The big cats and their fossil relatives: an illustrated guide to their evolution and natural history (Columbia University Press).
- Kurtén, B. (1968). Pleistocene mammals of Europe (Weidenfeld & Nicolson), pp. 1–317.
- Kurtén, B. (1972). The age of mammals (Columbia University Press), pp. 1–250.
- Rawn-Schatzinger, V. (1992). The scimitar cat, *Homotherium serum* cope: osteology, functional morphology, and predatory behavior. (Illinois State Museum Reports of Investigations), pp. 1–80.
- Widga, C., Fulton, T.L., Martin, L.D., and Shapiro, B. (2012). *Homotherium serum* and *Cervalces* from the Great Lakes Region, USA: geochronology, morphology and ancient DNA. *Boreas* *41*, 546–556.
- Barnett, R., Barnes, I., Phillips, M.J., Martin, L.D., Harington, C.R., Leonard, J.A., and Cooper, A. (2005). Evolution of the extinct Sabretooths and the American cheetah-like cat. *Curr. Biol.* *15*, R589–R590.
- Paijmans, J.L.A., Barnett, R., Gilbert, M.T.P., Zepeda-Mendoza, M.L., Reumer, J.W.F., de Vos, J., Zazula, G., Nagel, D., Baryshnikov, G.F., Leonard, J.A., et al. (2017). Evolutionary History of Saber-Toothed Cats Based on Ancient Mitogenomics. *Curr. Biol.* *27*, 3330–3336.e5.
- Westbury, M.V., Hartmann, S., Barlow, A., Preick, M., Ridush, B., Nagel, D., Rathgeber, T., Ziegler, R., Baryshnikov, G., Sheng, G., et al. (2020). Hyena paleogenomes reveal a complex evolutionary history of cross-continental gene flow between spotted and cave hyena. *Science Advances* *6*, <https://doi.org/10.1126/sciadv.aay0456>.
- Li, G., Figueiró, H.V., Eizirik, E., and Murphy, W.J. (2019). Recombination-Aware Phylogenomics Reveals the Structured Genomic Landscape of Hybridizing Cat Species. *Mol. Biol. Evol.* *36*, 2111–2126.
- Kurtén, B., and Anderson, E. (1980). Pleistocene mammals of North America (New York: Columbia University Press).
- Yamaguchi, N., Cooper, A., Werdelin, L., and Macdonald, D.W. (2004). Evolution of the mane and group-living in the lion (*Panthera leo*): a review. *J. Zool.* *263*, 329–342.
- Naples, V.L., Martin, L.D., and Babiary, J.P. (2011). The Other Saber-tooths: Scimitar-tooth Cats of the Western Hemisphere (JHU Press).
- Antón, M. (2013). Sabertooth (Indiana University Press).
- de Manuel, M., Barnett, R., Sandoval-Velasco, M., Yamaguchi, N., Garrett Vieira, F., Zepeda Mendoza, M.L., Liu, S., Martin, M.D., Sinding, M.S., Mak, S.S.T., et al. (2020). The evolutionary history of extinct and living lions. *Proc. Natl. Acad. Sci. USA* *117*, 10927–10934.
- Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* *24*, 1586–1591.
- Rothwell, T. (2003). Phylogenetic systematics of North American *Pseudaelurus* (Carnivora: Felidae). *American Museum Novitates* *2003*, 1–64.
- Gopalakrishnan, S., Sinding, M.S., Ramos-Madrugal, J., Niemann, J., Samaniego Castruita, J.A., Vieira, F.G., Carøe, C., Montero, M.M., Kuderna, L., Serres, A., et al. (2018). Interspecific Gene Flow Shaped the Evolution of the Genus *Canis*. *Curr. Biol.* *28*, 3441–3449.e5, <https://doi.org/10.1016/j.cub.2018.08.041>.
- Figueiró, H.V., Li, G., Trindade, F.J., Assis, J., Pais, F., Fernandes, G., Santos, S.H.D., Hughes, G.M., Komissarov, A., Antunes, A., et al. (2017). Genome-wide signatures of complex introgression and adaptive evolution in the big cats. *Sci. Adv.* *3*, e1700299.
- Li, G., Davis, B.W., Eizirik, E., and Murphy, W.J. (2016). Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). *Genome Res.* *26*, 1–11.
- Werdelin, L., and Sanders, W.J. (2010). Cenozoic Mammals of Africa (Univ of California Press).
- Louys, J. (2014). The large terrestrial carnivore guild in Quaternary Southeast Asia. *Quat. Sci. Rev.* *96*, 86–97.
- Antón, M., Galobart, A., and Turner, A. (2005). Co-existence of scimitar-toothed cats, lions and hominins in the European Pleistocene. Implications of the post-cranial anatomy of *Homotherium latidens* (Owen) for comparative palaeoecology. *Quat. Sci. Rev.* *24*, 1287–1301.
- Mirarab, S., and Warnow, T. (2015). ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* *31*, i44–i52.
- Sayyari, E., Whitfield, J.B., and Mirarab, S. (2017). DiscoVista: interpretable visualizations of gene tree discordance. arXiv, 1709.09305 [q-bio.PE]. <http://arxiv.org/abs/1709.09305>.
- Schaller, G.B. (2009). The Serengeti Lion: A Study of Predator-Prey Relations (University of Chicago Press).
- Mondal, M., Bertranpetit, J., and Lao, O. (2019). Approximate Bayesian computation with deep learning supports a third archaic introgression in Asia and Oceania. *Nat. Commun.* *10*, 246.
- Kuhlwiilm, M., Han, S., Sousa, V.C., Excoffier, L., and Marques-Bonet, T. (2019). Ancient admixture from an extinct ape lineage into bonobos. *Nat. Ecol. Evol.* *3*, 957–965.
- Stevens, E., Carss, K.J., Cirak, S., Foley, A.R., Torelli, S., Willer, T., Tambunan, D.E., Yau, S., Brodd, L., Sewry, C.A., et al.; UK10K Consortium (2013). Mutations in B3GALNT2 cause congenital muscular dystrophy and hypoglycosylation of α -dystroglycan. *Am. J. Hum. Genet.* *92*, 354–365.
- Patel, N., Aldahmesh, M.A., Alkuraya, H., Anazi, S., Alsharif, H., Khan, A.O., Sunker, A., Al-Mohsen, S., Abboud, E.B., Nowlaty, S.R., et al. (2016). Expanding the clinical, allelic, and locus heterogeneity of retinal dystrophies. *Genet. Med.* *18*, 554–562.
- Duong, H.A., Robles, M.S., Knutti, D., and Weitz, C.J. (2011). A molecular mechanism for circadian clock negative feedback. *Science* *332*, 1436–1439.
- Chiu, J.C., Ko, H.W., and Edery, I. (2011). NEMO/NLK phosphorylates PERIOD to initiate a time-delay phosphorylation circuit that sets circadian clock speed. *Cell* *145*, 357–370.
- Johansson, C., Willeit, M., Smedh, C., Ekholm, J., Paunio, T., Kieseppä, T., Lichtermann, D., Praschak-Rieder, N., Neumeister, A., Nilsson, L.-G., et al. (2003). Circadian clock-related polymorphisms in seasonal affective disorder and their relevance to diurnal preference. *Neuropsychopharmacology* *28*, 734–739.
- Turco, M., Biscontin, A., Corrias, M., Caccin, L., Bano, M., Chiaromanni, F., Salamanca, M., Mattei, D., Salvo, C., Mazzotta, G., et al. (2017). Diurnal preference, mood and the response to morning light in relation to polymorphisms in the human clock gene PER3. *Sci. Rep.* *7*, 6967.
- Casares-Hidalgo, C., Pérez-Ramos, A., Forner-Gumbau, M., Pastor, F.J., and Figueirido, B. (2019). Taking a look into the orbit of mammalian carnivores. *J. Anat.* *234*, 622–636.

36. Nordin, A., Larsson, E., Thornell, L.-E., and Holmberg, M. (2011). Tissue-specific splicing of ISCU results in a skeletal muscle phenotype in myopathy with lactic acidosis, while complete loss of ISCU results in early embryonic death in mice. *Hum. Genet.* *129*, 371–378.
37. Kim, S., Cho, Y.S., Kim, H.-M., Chung, O., Kim, H., Jho, S., Seomun, H., Kim, J., Bang, W.Y., Kim, C., et al. (2016). Comparison of carnivore, omnivore, and herbivore mammalian genomes with a new leopard assembly. *Genome Biol.* *17*, 211.
38. Anyonge, W. (1996). Locomotor behaviour in Plio-Pleistocene sabre-tooth cats: a biomechanical analysis. *J. Zool.* *238*, 395–413.
39. Lewis, M.E., and Werdelin, L. (2010). Carnivoran Dispersal Out of Africa During the Early Pleistocene: Relevance for Hominins? In *Out of Africa I: The First Hominin Colonization of Eurasia*, J.G. Fleagle, J.J. Shea, F.E. Grine, A.L. Baden, and R.E. Leakey, eds. (Dordrecht: Springer Netherlands), pp. 13–26.
40. Joly, K., Gurarie, E., Sorum, M.S., Kaczensky, P., Cameron, M.D., Jakes, A.F., Borg, B.L., Nandintsetseg, D., Hopcraft, J.G.C., Buuveibaatar, B., et al. (2019). Longest terrestrial migrations and movements around the world. *Sci. Rep.* *9*, 15333.
41. Abascal, F., Corvelo, A., Cruz, F., Villanueva-Cañás, J.L., Vlasova, A., Marcet-Houben, M., Martínez-Cruz, B., Cheng, J.Y., Prieto, P., Quesada, V., et al. (2016). Extreme genomic erosion after recurrent demographic bottlenecks in the highly endangered Iberian lynx. *Genome Biol.* *17*, 251.
42. Dobrynin, P., Liu, S., Tamazian, G., Xiong, Z., Yurchenko, A.A., Krashennikova, K., Kliver, S., Schmidt-Küntzel, A., Koepfli, K.-P., Johnson, W., et al. (2015). Genomic legacy of the African cheetah, *Acinonyx jubatus*. *Genome Biol.* *16*, 277.
43. Cho, Y.S., Hu, L., Hou, H., Lee, H., Xu, J., Kwon, S., Oh, S., Kim, H.-M., Jho, S., Kim, S., et al. (2013). The tiger genome and comparative analysis with lion and snow leopard genomes. *Nat. Commun.* *4*, 2433.
44. Westbury, M.V., Hartmann, S., Barlow, A., Wiesel, I., Leo, V., Welch, R., Parker, D.M., Sicks, F., Ludwig, A., Dalén, L., and Hofreiter, M. (2018). Extended and Continuous Decline in Effective Population Size Results in Low Genomic Diversity in the World's Rarest Hyena Species, the Brown Hyena. *Mol. Biol. Evol.* *35*, 1225–1237.
45. Schubert, M., Ermini, L., Der Sarkissian, C., Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M.D., Fernández, R., Kircher, M., McCue, M., et al. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nat. Protoc.* *9*, 1056–1082.
46. Schubert, M., Lindgreen, S., and Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes* *9*, 88.
47. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* *25*, 1754–1760.
48. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* *20*, 1297–1303.
49. Ginolhac, A., Rasmussen, M., Gilbert, M.T.P., Willerslev, E., and Orlando, L. (2011). mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* *27*, 2153–2155.
50. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* *29*, 1682–1684.
51. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. *Bioinformatics* *27*, 2156–2158.
52. Quinlan, A.R. (2014). BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr. Protoc. Bioinformatics* *47*, 11.12.1–34.
53. Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* *16*, 276–277.
54. Li, L., Stoeckert, C.J., Jr., and Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* *13*, 2178–2189.
55. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* *30*, 1312–1313.
56. Zhang, J., Nielsen, R., and Yang, Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.* *22*, 2472–2479.
57. Rambaut, A., Drummond, A.J., Xie, D., Baele, G., and Suchard, M.A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* *67*, 901–904.
58. Löytynoja, A., and Goldman, N. (2005). An algorithm for progressive multiple alignment of sequences with insertions. *Proc. Natl. Acad. Sci. USA* *102*, 10557–10562.
59. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* *25*, 2078–2079.
60. Korneliusson, T.S., Albrechtsen, A., and Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* *15*, 356.
61. Waters, M.R., Stafford, T.W., Jr., Kooyman, B., and Hills, L.V. (2015). Late Pleistocene horse and camel hunting at the southern margin of the ice-free corridor: reassessing the age of Wally's Beach, Canada. *Proc. Natl. Acad. Sci. USA* *112*, 4263–4267.
62. Stafford, T.W., Brendel, K., and Duhamel, R.C. (1988). Radiocarbon, ¹³C and ¹⁵N analysis of fossil bone: Removal of humates with XAD-2 resin. *Geochim. Cosmochim. Acta* *52*, 2257–2267.
63. Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., Schubert, M., Cappellini, E., Petersen, B., Moltke, I., et al. (2013). Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* *499*, 74–78.
64. Vilstrup, J.T., Seguin-Orlando, A., Stiller, M., Ginolhac, A., Raghavan, M., Nielsen, S.C.A., Weinstock, J., Froese, D., Vasiliev, S.K., Ovodov, N.D., et al. (2013). Mitochondrial phylogenomics of modern and ancient equids. *PLoS ONE* *8*, e55950.
65. Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* *2010*, <https://doi.org/10.1101/pdb.prot5448>.
66. dos Reis, M., and Yang, Z. (2011). Approximate likelihood calculation on a phylogeny for Bayesian estimation of divergence times. *Mol. Biol. Evol.* *28*, 2161–2172.
67. Van Valkenburgh, B., Grady, F., and Kurtén, B. (1990). The Plio-Pleistocene cheetah-like cat *Miracinonyx inexpectatus* of North America. *J. Vertebr. Paleontol.* *10*, 434–454.
68. Eizirik, E., Murphy, W.J., Koepfli, K.-P., Johnson, W.E., Dragoo, J.W., Wayne, R.K., and O'Brien, S.J. (2010). Pattern and timing of diversification of the mammalian order Carnivora inferred from multiple nuclear gene sequences. *Mol. Phylogenet. Evol.* *56*, 49–63.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological sample		
<i>Homotherium latidens</i>	This paper	Yukon Government Palaeontology Program
Lion (<i>Panthera leo</i>)	This paper	Copenhagen Zoo
Sand cat (<i>Felis margarita</i>)	This paper	Leibniz Institute for Zoo and Wildlife Research
Fishing cat (<i>Prionailurus viverrinus</i>)	This paper	Leibniz Institute for Zoo and Wildlife Research
Leopard cat (<i>P. bengalensis</i>)	This paper	Leibniz Institute for Zoo and Wildlife Research
Caracal (<i>Caracal caraca</i>)	This paper	Copenhagen Zoo
Chemicals, Peptides, and Recombinant Proteins		
AccuPrim Pfx DNA Polymerase	Invitrogen	Cat# 12344024
Phusion High-Fidelity DNA Polymerase	New England Biolabs	Cat# M0530S
KAPA HiFi HotStart polymerase	Roche	Cat# KK2801 07959052001
Critical Commercial Assays		
NEBNext DNA Library Prep Master Mix Set	New England Biolabs	Cat# E6070
MinElute PCR Purification Kit	QIAGEN	Cat# 28006
QIAquick column system	QIAGEN	Cat# 28104
myBaits target enrichment kit	Arbor Biosciences	NA
Kingfisher blood DNA extraction kit	Thermo Fisher Scientific	Cat# 98010196
PCR-free Truseq Illumina library kit	Illumina	Cat# 20015962
Deposited data		
<i>Felis nigripes</i>	99 Lives Cat Genome Sequencing Initiative	NCBI SRA accession code: SRR2511865
<i>Felis catus</i>	99 Lives Cat Genome Sequencing Initiative	NCBI SRA accession code: SRR2224864
<i>Felis chaus</i>	[20]	NCBI SRA accession code: SRR2062187
<i>Felis margarita</i>	This study	NCBI Bioproject accession code: PRJNA649575
<i>Felis margarita</i>	[20]	NCBI SRA accession code: SRR2062538
<i>Prionailurus viverrinus</i>	This study	NCBI Bioproject accession code: PRJNA649563
<i>Prionailurus bengalensis</i>	NA	NCBI SRA accession code: SRR2062628
<i>Prionailurus bengalensis</i>	This study	NCBI Bioproject accession code: PRJNA649572
<i>Lynx pardinus</i>	[41]	European nucleotide archive accession code: ERA562804
<i>Acinonyx jubatus</i>	[42]	NCBI SRA accession code: SRS1123638
<i>Caracal caracal</i>	This study	NCBI SRA sample accession code: SAMN15096300
<i>Neofelis nebulosa</i>	[15]	NCBI SRA sample accession code: SAMN14352199
<i>Panthera tigris altaica</i>	[43]	NCBI SRA accession code: SRR836311
<i>Panthera tigris tigris</i>	[43]	NCBI SRA accession code: SRR836354
<i>Panthera uncia</i>	[43]	NCBI SRA accession code: SRR836372
<i>Panthera onca</i>	[19]	NCBI SRA sample accession code: SAMN05907657
<i>Panthera pardus</i>	[37]	NCBI SRA accession code: SRR3041424
<i>Panthera leo</i>	[43]	NCBI SRA accession code: SRR836361
<i>Panthera leo</i>	[43]	NCBI SRA accession code: SRR836370
<i>Hyaena hyaena</i>	[44]	NCBI Bioproject accession code: PRJNA390068
<i>Crocuta crocuta</i>	[9]	NCBI Bioproject accession code: PRJNA554753
<i>Homotherium latidens</i>	This study	NCBI Bioproject accession code: PRJNA649760

(Continued on next page)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
Illumina compatible adapters	Illumina	NA
Hybridization capture probes	Arbor Biosciences	NA
Software and Algorithms		
PALEOMIX v1.2.5	[45]	https://github.com/MikkelSchubert/paleomix
AdapterRemoval v2.0.0	[46]	https://github.com/MikkelSchubert/adapterremoval
BWA v0.7.5a	[47]	http://bio-bwa.sourceforge.net/
Picard v2.18.0	NA	https://broadinstitute.github.io/picard
GATK v3.8.0	[48]	https://gatk.broadinstitute.org/hc/en-us
mapDamage v2.0.5	[49, 50]	https://ginolhac.github.io/mapDamage/
vcf-tools v0.1.14	[51]	https://vcftools.github.io/
bedtools v2.29.0	[52]	https://github.com/arq5x/bedtools2
EMBOSS v6.6.0	[53]	https://www.ebi.ac.uk/Tools/emboss/
OrthoMCL v2.0.9	[54]	https://orthomcl.org/
RAxML v8.2.11	[55]	https://sco.h-its.org/exelixis/software.html
ASTRAL-II	[24]	https://github.com/Smirarab/ASTRAL
DiscoVista	[25]	https://github.com/esayari/DiscoVista
PAML 4.5	[16, 56]	http://evomics.org/resources/software/molecular-evolution-software/paml/
Tracer	[57]	https://github.com/beast-dev/tracer/releases/latest
PRANK	[58]	https://www.ebi.ac.uk/research/goldman/software/prank
SAMtools v1.6	[59]	http://samtools.sourceforge.net/
ANGSD v0.921	[60]	https://github.com/ANGSD/angsd

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for reagents and data may be directed to and will be fulfilled by the Lead Contact, Michael V Westbury (m.westbury@sund.ku.dk).

Materials Availability

This study did not generate new unique reagents.

Data and Code Availability

NCBI accession codes for all newly generated raw fastq reads can be found in the key resources table.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Fossil information

Specimen YG 439.38 was sampled from the collections of the Yukon Government Palaeontology Program housed in Whitehorse, Yukon Territory, Canada, and was recovered from the placer gold mining site of Dominion Creek, near Dawson City, Yukon Territory. Initially identified as a large felid humerus by GZ, it was later assigned based on morphology to *Homotherium*, the scimitar-toothed cat. Here, we refer to our fossil material as *H. latidens* in accordance with the mitochondrial genomic data which suggested that all late Pleistocene Holarctic *Homotherium* belonged to a single species [8].

Radiocarbon dating

A section of bone was delivered to Stafford Research LLC for radiocarbon dating. Briefly, samples of crushed bone were decalcified and washed, treated with 0.05 N NaOH overnight to remove humics, soaked in 0.1 N HCl, gelatinised at 60°C at pH 2, and ultrafiltered at 30 kDa. Resultant collagen was then graphitised and run on an accelerator mass spectrometer according to standard protocol [61, 62].

METHOD DETAILS

Ancient DNA extraction, library preparation, and sequencing

Samples of cortical bone were taken from a long bone element (approx. 1 cm³) using a Dremel drill and reduced to powder in a Mikrodismembrator. DNA extraction was performed as described in Orlando et al. [63] in a dedicated ancient DNA laboratory in parallel with negative extraction controls. DNA extracts and negative controls were built into genomic libraries using the NEB E6070 kit and a slightly modified version of the protocol as used by Vilstrup et al. [64]. Briefly, the extract (30 µl) was end-repaired and passed through a MinElute column. The collected flow-through was adaptor-ligated and passed through a QiaQuick column. Adaptor fill-in reaction was performed on the flowthrough, before a final incubation at 37°C (30 min) followed by inactivation overnight at –20°C.

We amplified the resulting library in a 50 µl volume reaction, using 25 µl of library for 12 cycles under the following reaction conditions. Final concentrations of reagents in the master mix were: 1.25 U AccuPrime Pfx DNA Polymerase (Invitrogen), 1x AccuPrime Pfx reaction mix (Invitrogen), 0.4 mg/mL BSA, 120 nM primer inPE, 120 nM of a multiplexing indexing primer containing a unique 6-nucleotide index code (Illumina). PCR cycling conditions consisted of an initial denaturation step at 95°C for 2 min, followed by 12 cycles of 95°C denaturation for 15 s, 60°C annealing for 30 s, and 68°C extension for 30 s and a final extension step at 68°C for 7 min. Amplified libraries were checked for the presence of DNA on a 2% Agarose gel before purification using the QIAquick column system (QIAGEN) and quantification on an Agilent 2100 BioAnalyzer. Quantified libraries were communally pooled in equimolar ratios and sequenced as 100 bp single-end reads on three lanes of an Illumina HiSeq2000 platform at the Danish National High-Throughput Sequencing Centre and 100 bp paired-end reads on two lanes of Illumina HiSeq2000 at BGI Copenhagen.

Genome capture

Shotgun-sequenced libraries were assessed for endogenous content and the best libraries were taken forward for two sets of capture experiments. The first method used biotinylated RNA probes transcribed from fresh DNA extract derived from modern lion tissue in an attempt to capture the whole genome (whole genome capture/WGC). The second method used previously published lion genomic data [43] to identify exon coding regions and create biotinylated RNA baits that covered these regions (exome capture). Both types of baits were generated by Arbor Biosciences (Ann Arbor, MI, USA). Capture experiments using both sets of baits were carried out using the myBaits target enrichment kit and following the instructions described in manual V3 (<https://arborbiosci.com/wp-content/uploads/2017/10/MYbaits-manual-v3.pdf>) to enrich the ancient libraries for endogenous DNA. After capture and cleanup, enriched libraries were re-amplified for further sequencing using either Phusion polymerase or KAPA HiFi Hot-Start polymerase with primers IS5_reamp.P5 and IS6_reamp.P7 over 14 cycles [65]. Following quantification using a TapeStation instrument, amplified captured libraries were sequenced on an Illumina HiSeq2000 platform. In total we sequenced 303,621,590 single-end reads and 524,795,913 paired end reads.

Modern DNA extraction, library preparation, and sequencing

We extracted DNA from the sand cat (*Felis margarita*), fishing cat (*Prionailurus viverrinus*), leopard cat (*P. bengalensis*), and caracal (*Caracal caracal*) on a KingFisher Duo Prime robot using the Kingfisher blood DNA extraction kit (Thermo Fisher Scientific) according to the manufacturer's instructions. The extracts were built into PCR-free Truseq Illumina sequencing libraries and sequenced on an Illumina HiSeqX using 2 × 150 bp paired-end sequencing at the National Genomics Infrastructure (NGI) in Stockholm. The sand cat, fishing cat, leopard cat, and caracal had 1,706,953,684, 992,351,075, 1,264,528,157, 1,596,359,637 raw paired-end reads sequenced respectively.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data processing pipeline

Post-sequencing read processing of *Homotherium* was performed using the PALEOMIX v1.2.5 pipeline [45]. Removal of adaptor contamination and adaptor dimers, as well as trimming of low-quality bases (BaseQ < 5 or Ns) was accomplished with AdapterRemoval v2.0.0 [46], while also removing reads shorter than 30 bp or with over 50 bp missing data. All retained reads were mapped against the publicly available African lion reference genome [15] (NCBI Bioproject: PRJNA615082, NCBI Accession: JAAVKH000000000) with BWA-MEM v0.7.5a and default parameters [47]. We used Picard v2.18.0 (<http://broadinstitute.github.io/picard/>) to identify and filter PCR duplicates by the 5-prime end mapping coordinate. GATK v3.8.0 [48] was used to perform an insertion/deletion (indel) realignment step to adjust for increased error rates at the ends of short reads in the presence of indels. In the absence of a curated dataset of indels, this step relied on a set of indels identified in the specific sample being processed. DNA damage patterns were assessed, and base quality scores were recalibrated using mapDamage v2.0.5 [49, 50]. Genotypes were called with HaplotypeCaller from GATK v3.8. Data processing of the extant species followed the same pipeline with slight modifications: no minimum read length cut-off or missing data cut-off was applied during the adaptor trimming step, and bases were not recalibrated using mapDamage. These steps were removed as the data from the extant species was not highly fragmented and damaged as is expected for ancient DNA. For the sand cat, fishing cat, leopard cat, and caracal 1,661,576,497, 966,303,222, 1,229,116,793, 1,563,010,218 unique reads successfully mapped to the reference nuclear genome respectively. 195,372,434 *Homotherium latidens* unique reads successfully mapped to the nuclear genome and 14,900,381 successfully mapped to the exome.

Extracting coding and protein sequences

To obtain the coding sequence (CDS) regions for each species, we ignored all indels and calculated a consensus genome sequence using *vcf-consensus* from *vcf-tools* v0.1.14 [51]. Subsequently, since *vcf-consensus* assumes the reference sequence in the case of missing data, we masked all regions in which depth was lower than 5x using *maskFastaFromBed* from *bedtools* v2.29.0 [52]. Finally, we used the gene annotations (v2) from the reference African lion genome to extract the coding sequences from the masked genomic sequence, and converted them to protein using *transeq* from *EMBOSS* v6.6.0 [53]. Afterward, orthologs were identified using the *OrthoMCL* v2.0.9 pipeline [54] with a BLAST e-value cut-off of 10^{-5} and *mcl* inflation of 1.5, only retaining 1-1 orthologs for downstream analyses.

Phylogenetic analyses

Custom python scripts were used to construct a multiple sequence alignment of the genotypes for the CDS regions. To achieve this, diploid genotypes were extracted from the aforementioned VCF file for every position in the African lion reference genome with a CDS annotation (v2 annotation). Genotypes passing filters (minimum genotype quality of 20, minimum read depth of 4) were used to populate the multiple sequence alignments with IUPAC ambiguity codes, while genotypes not passing the filters were replaced by the character 'N'. Indel variants were ignored (replaced by 'N') such that the lengths of CDS regions were the same for every taxon. Genotypes from antisense CDS regions were reverse-complemented such that the multiple sequence alignment for each gene consisted of a concatenation of sense genotypes from all CDS regions.

For phylogenetic reconstruction, all of the aforementioned CDS alignments were concatenated into a supermatrix of 29,216,712 bp. Of these, 370,311 (1.3% of total) positions were completely undetermined. *RAXML* v8.2.11 [55] was used to perform a maximum-likelihood phylogenetic analysis, with 100 bootstrap replicates (Figure S1). The dataset was partitioned such that a separate GTR+G model was assigned to each of the three codon positions.

In addition to the phylogenetic tree estimated using the concatenated exon sequences, a separate phylogenetic tree was constructed for 13,164 genes, using the concatenated exon sequences for each gene. Similar to the supermatrix tree, a separate GTR+G model was assigned for the three codon positions. Using these maximum-likelihood gene trees estimated using *RAXML*, we inferred the species tree under the multispecies coalescent using *ASTRAL-II* [24]. Subsequently, we visualized the discordance between the 13,164 gene trees and the species tree. At each node, the local posterior probabilities were computed with *ASTRAL-II* [24], while the relative frequencies of the three possible bipartitions (possible arrangements of a quartet on an unrooted tree) induced by each internal branch of the estimated species tree were computed and visualized using *DiscoVista* v1.0 [25] (Figure S2).

Phylogenomic dating

We inferred the evolutionary timescale using a Bayesian phylogenomic dating analysis of the concatenated exon sequences from 20 species (Table S1). The sequence data were analyzed using *MCMCTree*, part of the *PAML* package [16]. Rate variation among lineages was modeled using a separate correlated-rates relaxed clock for each of the three codon positions. A uniform prior was specified for the relative node times. To evaluate the robustness of our date estimates, we repeated the analysis using an independent-rates relaxed clock model assigned to each codon position. We also varied the gamma prior on the mean substitution rate.

Approximate likelihood calculation was used to reduce computational expense [66]. Posterior distributions of node times were obtained by Markov chain Monte Carlo sampling. Following a burn-in of 10 million steps, we drew samples every 5,000 steps over a total of 100 million steps. To check for convergence to the stationary distribution, all analyses were carried out in duplicate. Sufficient sampling was checked by inspection of the trace files in *Tracer* v1.7.1 [57].

The phylogenomic dating analysis was calibrated using age constraints (uniform priors) on four divergence events: (i) 29–35 million years for the *Hyaenidae-Felidae* split, based on the reasoning described by Barnett et al. [7]; (ii) minimum age of 3.8 million years for the *Neofelis-Panthera* split, based on the oldest fossil of *Panthera* [2, 67]; (iii) minimum age of 1.5 million years for the split between snow leopard and tiger, based on the oldest fossil of *Panthera tigris* [2]; and (iv) minimum age of 5.3 million years for the split between *Lynx* and *Prionailurus+Felis*, based on the oldest fossil of *Lynx* [68].

Tests of positive selection

We constructed multiple sequence alignments of CDS sequences of 1-1 orthologs using *PRANK*, taking codon positions into account and default parameters [58]. The *CODEML* module in *PAML* 4.5 [16, 56] was used to calculate the d_N/d_S ratio (ω) with the one-ratio and free-ratio models. We used the one-ratio model to estimate the general selection pressure with a single d_N/d_S ratio across all branches. The free-ratio model was employed to investigate the d_N/d_S ratio of each branch. The branch-site model was additionally used to further investigate potential positive selection. We specified *Homotherium latidens* as the foreground species and *Panthera leo*, *Panthera pardus*, *Panthera onca*, *Panthera tigris tigris*, *Panthera uncia*, *Neofelis nebulosa*, *Acinonyx jubatus*, *Lynx pardinus*, *Felis catus*, and *Crocuta crocuta* as background species and *Hyaena hyaena* as outgroup for the tests. We assessed the statistical significance of the likelihood ratio with a 10% false discovery rate criterion (Table S2). Genes with a free-ratio > 2 were considered as highly significant for signatures of positive selection (Table S3).

Genetic diversity

To estimate the genetic diversity of our *Homotherium latidens* genome, we used two independent approaches to calculate heterozygosity. These included the exome-wide heterozygosity, and autosome-wide heterozygosity. For comparative purposes, we calculated the same metrics from single representatives from each of the 15 cat species included in our study.

To calculate the exome-wide heterozygosity, we first extracted only the coding regions from the mapped bam files using SAMtools v1.6 [59]. We ran the resultant exome only bam files through ANGSD v0.921 [60] for each individual independently, sampling allele frequencies (-doSaf 1), taking genotype likelihoods into account using the GATK algorithm (-GL 2 and -domajorminor 1), and specifying the following parameters; only include sites with at least 5x read depth (-mininddepth 5), only include reads with a mapping quality > 30 (-minmapq 30), only consider bases with a quality > 30 (-minq 30), only consider reads mapping to one location (-unique-only 1), adjust quality scores around indels (-baq 2). Moreover, we only considered transversions (-notrans 1) for all species as aDNA damage is known to cause C-T transitions (G-A on the reverse strand) which could artificially inflate the diversity estimates in the *Homotherium* relative to the less damaged extant species. From this we calculated a folded SFS to recover the number of sites that are either homozygous or heterozygous in each individual using the realSFS command as part of the ANGSD package, with a tolerance of 1e-8. To calculate the variance of these results we included the -nSites parameter specifying 200kb windows of covered bases.

Following this, we calculated the autosome-wide patterns of heterozygosity of the same individuals, excluding coding sequences. Coding regions were removed from the mapped bam files using bedtools v2.26.0 intersect [52]. To make results comparable between individuals, and due to the relatively low coverage of our *Homotherium latidens* individual, we first calculated the average read depth across the genome excluding coding regions of all individuals and downsampled all individuals to an average read depth of ~7.75x using SAMtools. To remove scaffolds putatively linked to sex chromosomes, we calculated the average read depth of the downsampled bam files for each scaffold in two individuals known to be male (African lion, and cheetah). Scaffolds that had an average read depth < 4.5x were deemed as candidates for originating from either the X or Y chromosome so were removed from further analyses. This resulted in the cumulative removal of ~50Mb from the downstream analyses. We used the same ANGSD protocol as for the exome but reduced the minimum read depth from 5x to 3x due to the much lower average read depth of the whole genome data and restricted the analysis to scaffolds > 1MB. Furthermore, as opposed to the 200kb windows of covered bases, to obtain the variance of heterozygosity across the autosomes, we calculated the heterozygosity for each scaffold independently.

Supplemental Information

Genomic Adaptations and Evolutionary

History of the Extinct Scimitar-Toothed Cat,

Homotherium latidens

Ross Barnett, Michael V. Westbury, Marcela Sandoval-Velasco, Filipe Garrett Vieira, Sungwon Jeon, Grant Zazula, Michael D. Martin, Simon Y.W. Ho, Niklas Mather, Shyam Gopalakrishnan, Jazmín Ramos-Madrigal, Marc de Manuel, M. Lisandra Zepeda-Mendoza, Agostinho Antunes, Aldo Carmona Baez, Binia De Cahsan, Greger Larson, Stephen J. O'Brien, Eduardo Eizirik, Warren E. Johnson, Klaus-Peter Koepfli, Andreas Wilting, Jörns Fickel, Love Dalén, Eline D. Lorenzen, Tomas Marques-Bonet, Anders J. Hansen, Guojie Zhang, Jong Bhak, Nobuyuki Yamaguchi, and M. Thomas P. Gilbert

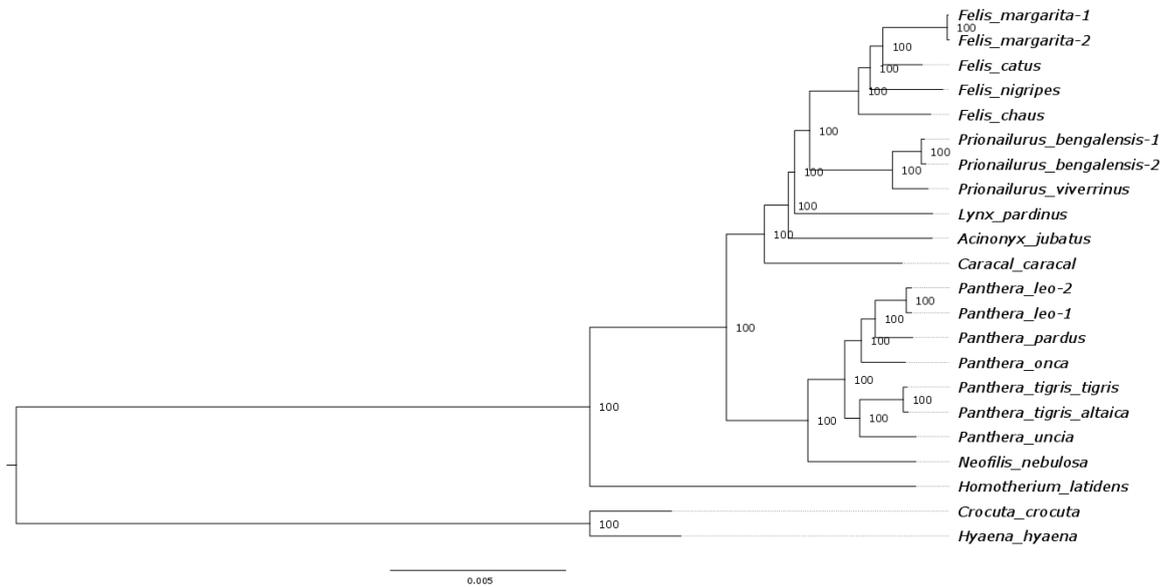
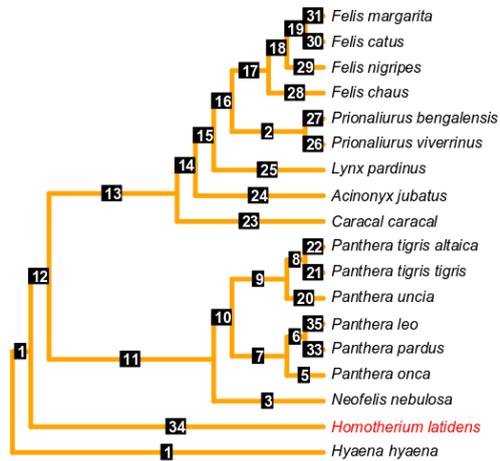


Figure S1: Maximum likelihood phylogenetic tree constructed using a concatenated supermatrix of 29,216,712 bp. Related to Figure 1 and STAR methods phylogenetic analyses. Numbers at the nodes show bootstrap support values. Branch lengths show the average number of substitutions per site.

A



B

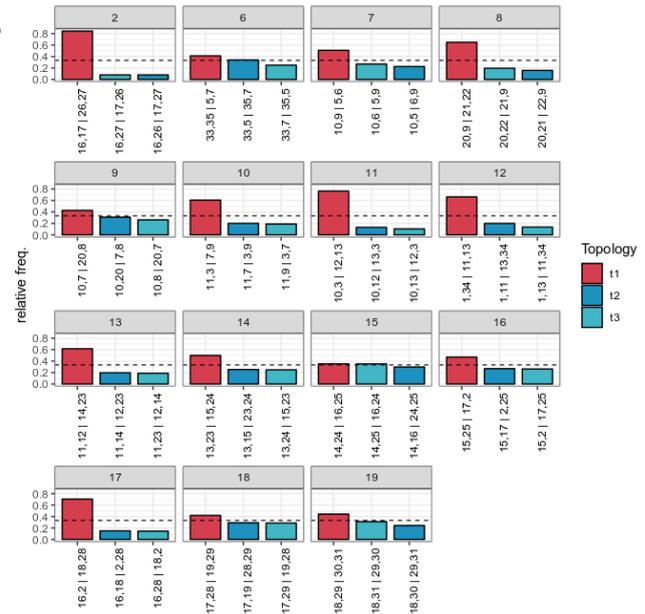


Figure S2. Phylogenetic analyses for tree topology discordances in the Felidae. Related to Figure 1 and STAR methods phylogenetic analyses. A) Species tree under the multispecies coalescent inferred from the maximum-likelihood gene trees. Branches have been numerically labelled for easy identification. B) Relative frequencies of the three possible bipartitions (possible arrangements of a quartet on an unrooted tree) induced by each internal branch of the estimated species tree. X axis numbers correspond to branch labels in A written in a quartet fashion based on the topology ([A,B],[C,D]). Dashed lines show the threshold value of $\frac{1}{3}$, shown theoretically to be the minimum frequency for a true bipartition. Related to Figure 1 and STAR methods phylogenetic analyses.

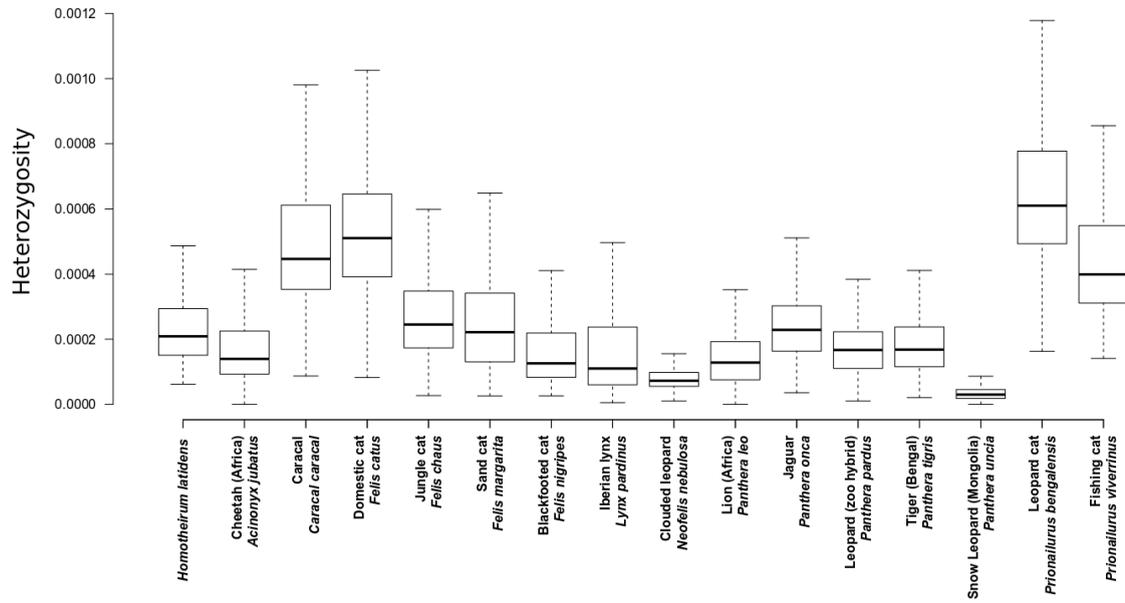


Figure S3: Exome-wide heterozygosity estimates based single representatives for each species used in the current study. Related to Figure 3 and STAR methods genetic diversity. Variance was estimated by independently calculating the average heterozygosity in non-overlapping windows of 200kb of covered bases.

Node	Node age (Ma) ^a				
	Correlated-rates model	Independent-rates model	Low rate prior: Gamma(1,50)	High rate prior: Gamma(1,2)	Prior (no data)
Root	32.8	32.3	33.0	32.8	32.4
<i>Homotherium</i> -sister	22.5	16.7	22.6	22.6	30.3
Pantherinae-Felinae	14.1	10.2	14.1	14.1	28.3
<i>Neofelis-Panthera</i>	8.3	5.5	8.3	8.3	23.3
<i>Panthera</i>	5.4	3.4	5.4	5.4	17.9
<i>Caracal</i> -sister	11.3	8.3	11.3	11.3	24.9
<i>Acinonyx</i> -sister	9.7	7.1	9.6	9.7	21.6
<i>Lynx</i> -sister	9.2	6.7	9.2	9.2	17.2
<i>Prionailurus-Felis</i>	8.2	6.0	8.1	8.2	12.7
<i>Felis</i>	5.0	3.8	5.0	5.0	9.2

Table S1 - Ages of key nodes in the phylogeny, estimated under a range of settings.

Related to Figure 1 and STAR methods phylogenomic dating. The time-tree in Figure 1 is based on the correlated-rates model. ^aCorrelated-rates model: Rates are assumed to be correlated between neighbouring branches in the phylogeny. Independent-rates model: Rates are assumed to be identically and independently distributed across branches in the phylogeny. Low rate prior: Gamma(1,50) prior for the rate, with a prior mean of 0.0002 substitutions/site/Myr. High rate prior: Gamma(1,2) prior for the rate, with a prior mean of 0.005 substitutions/site/Myr. Prior (no data): Analysis run without sequence data, such that the node ages are sampled from the joint prior.

GeneName	OneRatio	FreeRatio	Biological role	Functional role
AGBL5	0.97893	3.7966	vision sensitivity to diurnal light and circadian rhythms	vision
AK3	1	2.5265	endurance	mitochondria respiration
B3GALNT2	1	4.501	vision sensitivity to diurnal light and circadian rhythms	vision
C13orf30	1	2.3383	unknown	unknown
CAPNS2	1	2.8782	vision sensitivity to diurnal light and circadian rhythms	vision
ECSCR	1	3.4687	endurance	angiogenesis
F5	0.0001	3.3227	endurance	circulatory system
GCM1	1	2.0195	reproduction	placenta
GOLT1A	1	4.4857	cell function	Golgi
GPRC5A	1	2.7265	immunity	cancer
HMGB2	1	2.127	immunity	fecundity
IQCF5	0.64872	2.3567	unknown	unknown
ISCU	1	2.3704	endurance	mitochondria respiration
LAGE3	1	2.449	cell function	apoptosis
MIS18A	1	3.7168	cell function	mitosis
MMP12	1	2.2541	endurance	respiratory/circulatory system
NTF3	1	4.3076	socialisation	nervous system
OR11A1	0.83912	3.7452	olfaction	olfaction
Per1	0.02951	2.0949	vision sensitivity to diurnal light and circadian rhythms	circadian clock
PGD	1	2.8531	bone mineralization	bone mineralization
Rplp1	0.0001	3.1993	ribosome	protein synthesis
Rps13	1	2.4235	ribosome	protein synthesis
SCTR	1	2.8874	socialisation	social behaviour
SDPR	0.98745	2.3358	endurance	angiogenesis
SFPQ	1	2.4583	vision sensitivity to diurnal light and circadian rhythms	circadian clock

SLC1A7	1	2.1567	vision sensitivity to diurnal light and circadian rhythms	vision
SPACA3	0.0001	3.0126	reproduction	sperm
STAP1	0.99486	2.2366	cell function	docking protein
SURF1	1	2.1957	endurance	mitochondria respiration
TAF8	0.97884	2.9587	endurance	adipogenesis
TMEM45A	0.99662	4.899	endurance	hypoxia
unknown	0.99995	2.1152	unknown	unknown

Table S3: 31 genes under positive selection in the *Homotherium* genome with high values (free-ratio > 2) detected using one-ratio/free-ratio models and their respective hypothetical biological and functional roles. Related to Figure 2 and STAR methods and tests of positive selection.